

LEVERAGING VISION-CENTRIC GENERALIZABLE LEARNING FROM COARSE LOCATIONS AIRCRAFT OBJECT DETECTION BASED ON IMPROVED YOLO

Mrs. P. Yamuna¹, Akkala Bhanu Prakash², Ladineni Divija³, Radha Biradar⁴,
Sanem Ajay Kumar⁵

¹Assistant Professor, Computer Science and Engineering, ACE Engineering College, Ghatkesar,
Telangana, India.

^{2,3,4,5}Student, Computer Science and Engineering, ACE Engineering College, Ghatkesar, Telangana, India.

ABSTRACT

This abstract provides an overview of GNSS-based passive inverse synthetic aperture radar (SAR) imaging that uses global navigation satellite system (GNSS) signals to passively generate radar images of target areas or objects. The existing system is based on GNSS signals, emitted by satellite constellations like GPS are reflected off earth's surfaces and objects. The collected data is then processed to reconstruct SAR images, surface roughness and construct 2- D image. The proposed GNNS based P- ISAR algorithm not only focus and scale the ISAR image, but also provide an estimation of the cross- range velocity of the target. The utilization of global navigation satellite system (GNSS) signals for remote sensing has been a hot topic recently. This abstract provides an overview of GNSS-Based Passive Inverse Synthetic Aperture Radar (SAR) Imaging, that uses Global Navigation Satellite System (GNSS) signals to passively generate radar images of target areas or objects. The existing system is based on GNSS signals, emitted by satellite constellations like GPS are reflected off Earth's surfaces and objects. Specialized receivers capture these reflections as they move along predefined paths. The collected data is then processed to reconstruct SAR images, surface roughness and construct 2-D image. The proposed GNNS based P- ISAR algorithm not only focus and scale the ISAR image, but also provide an estimation of the cross- range velocity of the target and will help improve the target recognition capability of the GNSS based passive radar and also solves the problem of high cost signaling.

1. INTRODUCTION

Remote sensing is a contactless technique of gathering information. Since the beginning of earth observation from space, many satellites have launched into space, which has used successfully in a wide range of civil, agricultural, and military applications. Driven by the rapid development of remote sensing technologies, remote sensing images with finer resolution and clearer texture can be easily accessed by the modern airborne and space borne sensors. In response to the demand for automatic analysis of remote sensing data, object detection has been widely researched, among which aircraft detection occupies an important part owing to the successful application in the field of military and air transportation. However, aircraft detection remains a great challenge on account of the need for accurate and robust detection. Moreover, the complex background and noise, as well as the variation of spatial resolution of remote sensing images, increase the difficulty of this task.

Compared with passive sensors, such as optics and infrared, synthetic aperture radar (SAR) has the unique advantages of all-weather and day and night. It has outstanding strategic significance in military fields, such as battlefield situational awareness, typical target recognition, and precision guidance. Aircraft is an essential target in the civil field. The detection of it contributes to the effective management of airports. In the military field, the efficient and accurate acquisition of aircraft targets in the airport and airspace is of great importance, which can help to acquire battlefield military information and make battle plans in real time. Therefore, detecting aircraft targets based on SAR images is a significant research direction.

Unlike optical images, SAR imaging has a longer wavelength, a more complex imaging mechanism, and a more difficult visual interpretation of the imaging results. Therefore, SAR aircraft target detection faces some challenges. First, the target in the SAR image is discontinuous, which is composed of multiple discrete irregular scattering center bright spots. It is not easy to detect the complete aircraft target in this case. In addition, there are significant differences in target scales and many weak and small targets in SAR images, which makes detection difficult, which will be confused with aircraft target components to a certain extent, making it challenging to locate and identify the aircraft accurately. The last decade has seen a huge increase of available high resolution satellite images, which are used more and more for surveillance tasks. When monitoring military sites, it is necessary to automatically detect and identify objects of interest to derive trends. In this domain, aircraft recognition is of particular interest: each aircraft model has its own role, and a variation in the number of a specific type of aircraft at a given location can be a highly

relevant insight. This recognition task needs to be reliable to allow the automation of site analysis in particular to derive alerts corresponding to unusual events. Robustness to noise, shadows, illumination or ground texture variation is challenging to obtain but mandatory for real-life applications.

With the continuous development of satellite remote sensing technology, the information amount of high resolution remote sensing images has increased sharply, and the detailed information contained in is getting more abundant. Some sensitive targets such as ships, tanks, airplanes and ports can also be clearly visible to naked eyes, for which the detection methods have become a hot spot for scholars. Aircraft play an irreplaceable role in both the civilian and military fields. Therefore, the detection method of aircraft targets in remote sensing images is of great significance. However, the detection of aircraft targets in remote sensing images remains to be a challenging problem because it is susceptible to interference of external factors such as weather, light, shadows, etc. Besides, when there are small targets in the images with high exposure and complex background, the difficulty of aircraft detection is expected to rise.

Many solutions have been proposed to solve the above problems of target detection. Traditional methods such as template matching are fast, simple and easy to implement, but they have high requirements on the target state and target size and perform badly in complex backgrounds. The machine learning methods are designed to be flexible and highly targeted, but they are solidified and have poor robustness. In recent years, deep learning methods have developed rapidly. Many target detection algorithms based on CNN (Convolutional Neural Networks) have been proposed and applied to target detection in remote sensing images. At present, target detection methods can be classified into two main types: Two-Stage methods and One-Stage methods.

The Two-Stage method is a deep convolutional network based on the candidate region. It first generates possible candidate blocks containing the detection target, and then classify and correct the candidate blocks and obtain the detection frame to achieve target detection. The more common algorithms are R-CNN, Fast R-CNN and Faster R-CNN etc. These methods have high detection accuracy, but low speed. The One-Stage method is based on the target detection of the deep convolutional network of regression calculation, which uses an end-to-end target detection method, such as SSD (Single Shot Detector), YOLO series and so on. These methods have a faster detection speed and can meet real-time requirements.

2. METHODOLOGY

1. Data Collection and Preprocessing

1. Data Collection:

- **Source:** Acquire high-resolution SAR images from publicly available datasets or proprietary sources such as military satellites, commercial SAR satellite providers, and airborne SAR systems.
- **Labeling:** Ensure that each image is annotated with the location and type of aircraft. Utilize tools like LabelImg or custom annotation tools for accurate labeling.

2. Data Preprocessing:

- **Normalization:** Normalize the pixel values of the SAR images to a standard range to facilitate better learning by neural networks.
- **Augmentation:** Apply data augmentation techniques such as rotation, scaling, translation, and flipping to increase the diversity of the training data and improve the robustness of the model.
- **Noise Reduction:** Implement noise reduction techniques specific to SAR images, such as filtering and despeckling, to enhance the quality of the images.

2. Model Selection

1. Model Architecture:

- **Two-Stage Methods:** Select models such as Faster R-CNN for their high accuracy in detecting aircraft in complex backgrounds. These models first generate region proposals and then refine these proposals to detect aircraft.
- **One-Stage Methods:** Choose models like YOLOv4 or SSD for their real-time detection capabilities. These models directly predict bounding boxes and class probabilities in a single step.

2. Feature Extraction:

- Use convolutional neural networks (CNNs) as the backbone for feature extraction. Popular choices include ResNet, VGG, and Efficient Net due to their proven effectiveness in capturing hierarchical features from images.

3. Training

1. Dataset Splitting:

- Split the dataset into training, validation, and test sets with a typical ratio of 70:20:10 to ensure the model is trained, validated, and tested on different subsets of data.

2. Hyperparameter Tuning:

- Tune hyperparameters such as learning rate, batch size, number of epochs, and network depth using techniques like grid search or Bayesian optimization to find the optimal settings for the model.

3. Training Process:

- Train the model using the training dataset while periodically validating it using the validation dataset to monitor performance and avoid overfitting.
- Implement early stopping and learning rate scheduling to enhance the training process.

4. Loss Function:

- Use a combination of classification loss (e.g., cross-entropy loss) and localization loss (e.g., smooth L1 loss) to train the model to accurately classify and localize aircraft in the images.

4. Evaluation

1. Performance Metrics:

- Evaluate the model using metrics such as Precision, Recall, F1 Score, and Mean Average Precision (mAP) to assess the detection accuracy and robustness.
- Perform detailed error analysis to understand common failure modes and areas for improvement.

2. Validation and Testing:

- Validate the model on the validation set and fine-tune based on performance.
- Test the final model on the test set to evaluate its generalization capability.

3. MODELING AND ANALYSIS

1. IMAGE PROCESSING

The aim of pre-processing is to improve the quality of the image so that we can analyze it in a better way. By preprocessing we can suppress undesired distortions and enhance some features which are necessary for the particular application we are working for. Those features might vary for different applications.

Image preprocessing is the steps taken to format images before they are used by model training and inference. This includes, but is not limited to, resizing, orienting, and color corrections.

Every image is made of pixels. And each pixel will have some intensity. Based on the intensity we can say if it is a white pixel or black pixel or something in between them. A histogram of an image is the representation of the intensity versus the number of pixels with that intensity. For example, a dark image will have many pixels which are black and few which are white. Representing that like a graph is what is called a histogram.

2. MODEL TRAINING

In the YOLO architecture we are using there are multiple output layers giving out predictions. First, the image is divided into various grids. Each grid has a dimension of $S \times S$. The following image shows how an input image is divided into grids. YOLO uses a single bounding box regression to predict the height, width, center, and class of objects.

In the image above, represents the probability of an object appearing in the bounding box. Finally we look at the detections that are left and draw bounding boxes around them and display the output image. Intersection over union (IOU) is a phenomenon in object detection that describes how boxes overlap. YOLO uses IOU to provide an output box that surrounds the objects perfectly. Each grid cell is responsible for predicting the bounding boxes and their confidence scores. The IOU is equal to 1 if the predicted bounding box is the same as the real box. This mechanism eliminates bounding boxes that are not equal to the real box

3. MODEL EVALUATION

The Machine Learning (ML) Module leverages advanced algorithms to analyze user financial data. ML models enhance the accuracy of expense categorization and can predict potential budget deviations based on historical data, contributing to more informed financial decision-making.

Intersection Over Union (IOU) is measure based on Jaccard Index that evaluates the overlap between two bounding boxes. It requires a ground truth bounding box. By applying the IOU we can tell if a detection is valid (True Positive) or not (False Positive). IOU is the ratio of the intersection of the of predicted bounding box and ground truth bounding box to the union of predicted bounding box and ground truth areas.

Concepts used by the metrics:

True Positive (TP): A correct detection. Detection with IOU threshold False Positive (FP): A wrong detection, Detection with IOU < threshold

False Negative (FN): A ground truth not detected, [if IOU with ground truththreshold.

4. RESULTS AND DISCUSSION

a. Graphs:

The graph illustrates the distribution of various activities among individuals, showcasing the engagement levels in activities such as walking, standing, descending stairs, ascending stairs, and lying down. The graph illustrates the distribution of activities across each individual, where each person is represented along the x-axis, and the y-axis indicates the count of activities performed by each person.

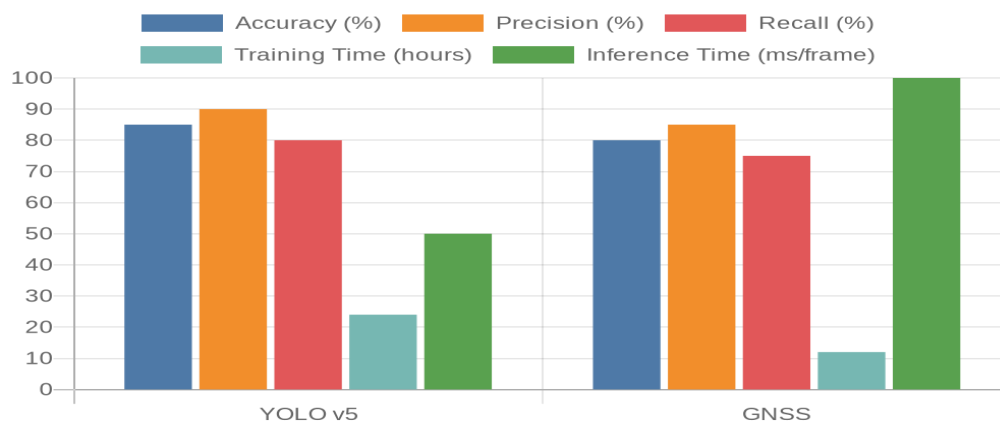


Figure 3: comparison between YOLO and GNSS

id	image_id	geometry	class	bounds	width	height
0	1 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(135, 522), (245, 522), (245, 600), (135, 600...	Airplane	(135, 522, 245, 600)	110	78
1	2 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(1025, 284), (1125, 284), (1125, 384), (1025,...	Airplane	(1025, 284, 1125, 384)	100	100
2	3 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(1058, 1503), (1130, 1503), (1130, 1568), (10...	Airplane	(1058, 1503, 1130, 1568)	72	65
3	4 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(813, 1518), (885, 1518), (885, 1604), (813, ...	Airplane	(813, 1518, 885, 1604)	72	86
4	5 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(594, 938), (657, 938), (657, 1012), (594, 10...	Airplane	(594, 938, 657, 1012)	63	74
5	6 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(451, 725), (524, 725), (524, 798), (451, 798...	Airplane	(451, 725, 524, 798)	73	73
6	7 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(1543, 1437), (1614, 1437), (1614, 1497), (15...	Airplane	(1543, 1437, 1614, 1497)	71	60
7	8 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(1485, 1370), (1561, 1370), (1561, 1437), (14...	Airplane	(1485, 1370, 1561, 1437)	76	67
8	9 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(1968, 1624), (2053, 1624), (2053, 1687), (19...	Airplane	(1968, 1624, 2053, 1687)	85	63
9	10 4f833867-273e-4d73-8bc3-cb2d9ceb54ef.jpg	[(1659, 1674), (1733, 1674), (1733, 1746), (16...	Airplane	(1659, 1674, 1733, 1746)	74	72

Figure 4. Calculated values of the parameters (Height, Width and Velocity)

b. Mathematical Formula:

1.Precision (P):

- Precision measures the accuracy of the model's positive predictions.
- It is calculated as the ratio of true positive (TP) detections to the total number of positive detections (TP + False Positives (FP)).
- $P = \frac{TP}{TP + FP}$

2. Recall (R):

- Recall measures the completeness of the model's positive predictions.
- It is calculated as the ratio of true positive (TP) detections to the total number of actual positive instances (TP + False Negatives (FN)).

$$\diamond R = TPTP + FNR = TP + FNTF$$

3. F1 Score:

- F1 Score is the harmonic mean of precision and recall.
- It provides a balance between precision and recall, considering both false positives and false negatives.
- F1 Score is calculated using the following formula:

$$\diamond F1 = 2 \times P \times R / (P + R)$$

4. Confusion Matrix:

- The confusion matrix provides a tabular representation of the model's predictions against ground truth values.
- It consists of four components: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).
- Each cell in the matrix represents the count of predictions falling into these categories.
- Here's a summary of the parameters and their formulas for performance evaluation:
- Precision (P): $P = TP / (TP + FP)$
- Recall (R): $R = TP / (TP + FN)$
- F1 Score: $F1 = 2 \times P \times R / (P + R)$
- Confusion Matrix:

Predicted Negative	Predicted Positive	Actual Negative	TP	Actual Positive	FN	TP	Actual Negative	Actual	Positive
Predicted Negative	TN	FN	Predicted Positive	FP	TP				

These parameters provide a comprehensive evaluation of the model's performance in terms of accuracy, completeness, and classification capabilities.

5. Average Precision (AP):

- Average Precision computes the area under the precision-recall curve.
- It provides a single scalar value that summarizes the precision-recall trade-off.
- AP is calculated by taking the average precision at each recall value across all classes.
- The mean Average Precision (mAP) is often used when evaluating object detection models across multiple classes.

The Average Precision is calculated as follows:

1. Calculate precision and recall values for different confidence thresholds.
2. Plot the precision-recall curve.
3. Compute the area under the precision-recall curve (integral of precision with respect to recall)

6. Mean Average Precision (mAP):

- mAP is the average of AP values across all classes.
- It provides a consolidated measure of the model's performance across all classes.

In object detection tasks, mAP is a widely used metric as it accounts for both precision and recall across various confidence thresholds.

Adding Average Precision and Mean Average Precision to the performance evaluation provides a more comprehensive understanding of the model's detection capabilities across different classes and confidence levels.

5. CONCLUSION

In summary, this system is a notable advancement in computer vision, particularly in object detection and velocity estimation tasks. Its lightweight framework emphasizes efficiency and adaptability for real-world applications. Techniques like random clipping and adaptive anchor frame calculation enhance dataset richness and prediction accuracy, while dropout layers ensure robustness against data variations. The system's efficiency, achieved through quantization and deployment optimizations, makes it practical for resource-constrained environments.

The integration of adaptive anchor frame calculation and dropout layers significantly improves prediction accuracy and network resilience. These features enable stable inference even in noisy or changing data scenarios. Moreover, the system's efficiency optimizations ensure scalability and accessibility, making it suitable for deployment in various domains requiring reliable object detection and velocity estimation.

6. REFERENCES

- [1] CGC-NET: Aircraft Detection in Remote Sensing Images Based on Lightweight Convolutional Neural Network - Ting Wang , Xiao dong Zeng, Chang qing Cao, Wei Li Zhejun Feng, Jin Wu, Xu Yan and Zeng yan Wu-2022-<https://ieeexplore.ieee.org/document/9736640>
- [2] Aircraft Target Detection in Remote Sensing Images Based on Improved YOLOv5 ShunLuo,JuanYu,YunjiangXiandXiaoLiao-2022
https://www.researchgate.net/publication/357640247_Aircraft_Target_Detection_in_Remote_Sensing_Images_Based_on_Improved_YOLOv5
- [3] SEFEPNet: Scale Expansion and Feature Enhancement Pyramid Network for SAR Aircraft Detection With Small Sample Dataset - Peng Zhang, Hao Xu, Tian Tian, Peng Gao, Linfeng Ti, Nan Zhang and Jin wen Tian-2022-<https://ieeexplore.ieee.org/document/9761751>
- [4] Multivariate Combined Collision Detection for Multi-Unmanned Aircraft Systems- Hon ghai Zhang , Jin peng Zhang , Gang Zhong , Hao Liu and Wen quan Liu- 2022- <https://ieeexplore.ieee.org/document/9904597>
- [5] An Experimental Test bed for the Study of Visual Based Navigation Docking of Two Vertical Compound Aircraft - Dong Wang , Qizhen Hong, Jing Wang, Heran Sun , Luchao Cheng , Mingyang Li, Congjing Wang, Xin Huang, Zhiyuan Wang and Jiahang Li - April 2021 – <https://ieeexplore.ieee.org/document/9433599>
- [6] A Multimode Anomaly Detection Method Based on OC-ELM for Aircraft Engine System- Shaowei Chen , Meng Wu, Pengfei We , Fangda Xu, Shengyue Wang and Shuai Zhao- 2021 - <https://ieeexplore.ieee.org/document/9349499>
- [7] Inspection interval optimization for aircraft composite structures with dent and delamination damage - Cai Jing and Dai Dingqiang -2021<https://ieeexplore.ieee.org/document/9369167>
- [8] Convolutional Neural Network Based Weakly Supervised Learning for Aircraft Detection From Remote Sensing Image- Zhi-Ze Wu , Thomas Weise , Yan Wang and Yongjun Wang- 2019- <https://ieeexplore.ieee.org/document/9178761>
- [9] A Novel Data Augmentation Method for Detection of Specific Aircraft in Remote Sensing RGB Images-Yiming Yan , Yumo Zhang and Nan Su-2019- <https://ieeexplore.ieee.org/document/8698795>
- [10] Real-Time FPGA-Based Hardware Neural Network for Fault Detection and Isolation in More Electric Aircraft- Qin Liu , Tian Liang , Zhen Huang and Venkata Dinavah-2019- <https://ieeexplore.ieee.org/document/888974>